outhwest

Great Plains

Research Note RM-527

February 1994

USDA Forest Service

Rocky Mountain Forest and Range Experiment Station

Using Conditional Utility Models for Measuring Welfare

Robert Mendelsohn[1]
Rosa Matzkin[1]
George Peterson[2]
Donald Rosenthal[3]

This paper argues that the functional form of conditional indirect utility functions is more restrictive than practitioners realize. The conditional indirect utility must be a function of the difference between income and price. Many empirical examples of discrete choice violate this constraint, implicitly assuming that the marginal utility of income in the conditional indirect utility function is constant across choices. When applied to conditional utility functions, this is equivalent to assuming that people of different incomes would make the same choices, which is frequently not the case. Empirical discrete choice studies should be interpreted using unconditional utility.

Keywords: Welfare measurement, conditional utility, discrete choice, marginal utility of income, unconditional utility

Introduction

In recent years, logit and probit models have grown in popularity as a tool for analyzing discrete choices because of their analytic simplicity and powerful empirical properties. In order to interpret these models, McFadden (1981), followed by Small and Rosen (1981) and Hanemann (1984), have proposed the use of random conditional utility models. Although this analytical development has been flawless, its adoption by the empirical literature has been more problematic. Empirical analysts appear not to understand the restrictions that must be placed on the form of conditional utility functions to make them consistent with theory. By using functional forms that have no utility basis or that require strong underlying assumptions, the welfare implications of this research is jeopardized. Given the widespread adoption of discrete choice models for modeling contingent valuation responses (Bishop and Heberlein 1979; Bishop et al. 1983; Bowker and Stoll 1981; Boyle and Bishop 1985; Loehman and De 1982; Sellar et al. 1985, 1986) and recreation decisions (Caulkins et al. 1986, Kling 1987, Morey and Rowe 1985, Smith and Kaoru 1986) it is important that researchers conducting future empirical work understand the restrictions theory places on the form of conditional utility functions.

This paper begins by clarifying the limitations of using conditional utility models for predicting welfare when consumers must make discrete (1,0) choices. Following McFadden, conditional indirect utility functions are deduced from an underlying direct utility function. The correct specification of the conditional utility model for discrete choices is emphasized. Although the architects of this methodology may argue that there is no need to clarify the theory, the applied literature appears not to understand the limitations of the model. Unless these methodological problems are openly reviewed, there is every reason to suspect empirical studies will continue to be flawed by suspect specifications.

[1] *Mendelsohn and Matzkin are Professors, Yale University, New Haven, Connecticut.*

[2] *General Research Supervisory Engineer, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado. Headquarters is in Fort Collins in cooperation with Colorado State University.*

[3] *Senior Economist, U.S. Department of Energy, Washington, D.C.*

In light of this theoretical discussion, several applications of discrete choice from the resource valuation literature are reviewed. The specifications used in many of these studies are analyzed for their theoretical consistency. It is shown that a few studies are correctly specified, but most are either correct only under strict assumptions or are simply misspecified.

Specifying Conditional Utility Differences

The general model we are analyzing is a discrete choice by consumers either to make or not to make a single purchase from one or more mutually exclusive options. This choice is assumed to be motivated by an underlying conditional indirect utility function, $V_i(Y,P_i,Z_i,e_i)$, associated with choosing good i, where Y is income, P_i is the price of good i, Z_i is a vector of characteristics associated with choice i, and e_i is an error term. The error term could be viewed as known to the decision-maker or a random variable. Since this distinction is not relevant to our argument, we assume for simplicity of exposition that the error term is known to the decision-maker. Each choice is presumed to have a conditional utility function associated with it. The model has the general form:

If $V(Y, P_i, Z_i, e_i) > V(Y, P_k, Z_k, e_k)$, for all k not equal to i, then good i is chosen; otherwise good i is not chosen.

In the single good example $V_k(Y, P_k, Z_k, e_k) = V_0(Y, 0, 0, e_0)$.

We derive, following McFadden (1981), the conditional indirect utility function from a general utility function in order to clarify the restrictions on its functional form. We focus entirely upon a pure (1,0) discrete choice. We begin with a continuous direct utility function $U(X,Z,e)$ defined over a set of goods X where e is a vector of individual characteristics known to the consumer but not to the econometrician. Thus, we assume that the consumer makes choices with certainty. The consumer faces the problem of choosing a bundle of goods to maximize his/her utility function subject to a budget constraint:

$$MAX\ U(X_1, X_2, \ldots X_N, X_{N+1}, e) \quad \text{s.t. } Y = \sum_{i=1}^{N} P_i X_i + X_{N+1} \qquad [1]$$

where prices have been normalized so that the price of the composite good X_{N+1} is equal to 1. The discrete choice model takes a snapshot of this decision process during which the consumer chooses only one unit of one good from a set of N goods. Given this constraint, one can solve for X_{N+1} in the budget constraint:

$$X_{N+1} = Y - P_i$$

where P_i is the price of the good that is selected.[4] Substituting the equation for X_{N+1} back into [1] yields:

$$\begin{aligned} &MAX\ U(1,0,0\ldots 0, Y-P_1, Z_1, e) \\ &\text{or } U(0,1,0,\ldots 0, Y-P_2, Z_2, e) \\ &\text{or } U(0,0,\ldots 0,1, Y-P_N, Z_N, e). \end{aligned} \qquad [2]$$

Each element above is a conditional indirect utility function. The conditional indirect utility function for a discrete choice consequently has the following form:

$$V(Y, P_i, Z_i, e_i) = V(Y - P_i, Z_i, e_i). \qquad [3]$$

The essential aspect of [3] that we wish to draw attention to is that neither income nor price enters the conditional indirect utility function alone. The price-income variable in the conditional utility function is the difference between income and price.

For example, suppose the indirect utility function [3] is linear:

$$V_i = A_i + B_i (Y - P_i) + C_i Z_i + e_i. \qquad [4]$$

In this case, the income and price coefficient must be the same magnitude and opposite sign for the discrete choice model to be consistent with the underlying direct utility function. Alternatively, one could assume that [3] has a semi-log functional form:

$$V_i = A_i + B_i \log (Y - P_i) + C_i \log Z_i + e_i. \qquad [5]$$

Regardless of the functional form, the difference between income and price is the appropriate argument (variable) in all conditional indirect utility functions.

As discussed above, the decision calculus is based upon the difference between the conditional indirect utility functions of each choice. For example, in the linear case, where the choice is to purchase the good at price P or not, the utility difference is:

$$\begin{aligned} \Delta V = V(Y,0) - V(Y-P,Z) = A_0 - A_1 + \\ B_0 Y - B_1 (Y-P) + C_1 Z + u \end{aligned} \qquad [6]$$

where $u = e_0 - e_1$. Similarly, with the semi-log indirect utility function [5], the utility difference is:

$$\begin{aligned} \Delta V = V(Y,0) - V(Y-P,Z) = A_0 - A_1 + \\ B_0 \log Y - B_1 \log (Y-P) - C_1 \log Z + u. \end{aligned} \qquad [7]$$

Of all the empirical papers in the applied resource economic literature, only Bowker and Stoll (1981)

[4] *Many applications of discrete choice in environmental/resource economics involve preferences for public goods for which there may be no market price. In the case of outdoor recreation, the effective price, however, may be transportation costs to the site. Further, in contingent valuation exercises, the price is the payment being requested of the individual on behalf of the public good.*

and Smith and Kaoru (1986) choose a functional form that is clearly consistent with [6] or [7].

In contrast, many empirical studies have introduced the following simplification of these two models:

$$\Delta V = V(Y,0) - V(Y-P,Z) = A_0 - A_1 + B_1 P + C_1 Z + u \quad [8]$$

and

$$\Delta V = V(Y,0) - V(Y-P,Z) = A_0 - A_1 + B_1 \log(P) + C_1 \log Z + u. \quad [9]$$

For example, Bowker and Stoll (1981), Caulkins et al. (1986), Kling (1987), Morey and Rowe (1985), and Sellar et al. (1985) all adopt [8]. Bishop and Heberlein (1979), Bishop et al. (1983), Boyle and Bishop (1985), Hanemann (1984), and Sellar et al. (1986) all adopt [9]. It is not possible to derive [9] from [7] unless Y approaches $2P$.

With [8], the literature has either explicitly or implicitly assumed that $B_0 = B_1$. Is this a reasonable assumption? Since B_i is the marginal utility of income and it is unlikely that the marginal utility of income will change depending upon the decision to consume one good, several authors argue that, in most circumstances, it is a reasonable assumption.

In this paper, we argue that $B_i \neq B_k$ in most circumstances. B_i is the marginal utility of income for a conditional utility function, not an unconditional utility function. If B_i was the marginal utility of income for an unconditional utility function, the argument for equality would be reasonable. However, the fact that B_i is the marginal utility of income for a conditional utility function is very important. Assuming that the conditional marginal utility of income is the same for all choices is effectively assuming that income will not affect the choice between the available alternatives. That is, people with high income and people with low income will choose alternative i with the same probability.[5] In the context of a traditional demand function, assuming income will not affect choice is equivalent to assuming that the income elasticity of the good is zero. Alternatively, assuming that $B_i = B_k$ for a class of goods is the same as assuming that the income elasticity of all goods k and i is the same.

Although there are special cases where income will not affect choice, the frequency by which people choose any environmental or recreation good i is often affected by their incomes. Empirical evidence from multiple site analyses reveals that B_i is not the same across sites. For example, in the three sites in Smith and Kaoru (1986), B_i varies from –.37 to –3.6, a statistically significant difference. Multiple site, travel cost analyses such as Burt and Brewer (1971) and Cichetti et al. (1976) also reveal that B_i will vary across sites. Individual travel cost analyses of sites often reveal that income elasticities are nonzero. Empirically, income does affect public recreation and environmental choices. In order to reflect the impact of income in a conditional utility function framework, one must assume that $B_i \neq B_k$ even for relatively inexpensive goods. In which case, the utility difference model must be specified as [6] or [7], not [8] or [9].

One final specification in the literature is provided by Bowker and Stoll (1981) and Loehman and De (1982), who estimate a log model:

$$\Delta V = V_i - V_0 = A_i + B_i \log Y + C_i \log P_i. \quad [10]$$

This model follows from an underlying indirect utility function of

$$V_i = A_i + B_i \log(Y / P_i). \quad [11]$$

However, [11] is not consistent with [3], and [10] is a poor approximation of [7].

In this paper, we argue against interpreting the results of models such as [8], [9], and [10] using conditional utility models. Because the specifications are inconsistent with the theory, the interpretations will be biased. For example, Bowker and Stoll (1981) reveal that the misspecified equations (relative to the correctly specified equation) can lead to biases in predicted welfare of between 25% to 100% of the true value.

However, we are not arguing that models such as [8], [9], and [10] are unattractive econometric models. In fact, these models perform admirably in empirical applications. Bowker and Stoll (1981), Boyle and Bishop (1985), and Sellar, Chavas, and Stoll (1986) all find that theoretically incorrect functional forms fit the data more closely than tested theoretically correct specifications. Given that [8], [9], and [10] can all be justified using a standard unconditional utility framework, the empirical results suggest they are preferable functional forms. The results suggest that conditional utility functions are not useful frameworks for analyzing discrete choices because they impose restrictions on the data that are not justified.

Conclusion

This paper examines the derivation of discrete choice (1,0) conditional indirect utility functions. The derivation implies that the conditional utility function is a function of the difference between income and price. Simplified functional forms that just depend upon price can only be used if income does not affect choice. The empirical valuation literature, however, reveals that income frequently does affect choice. Further, valuation exercises that compare correctly specified and incorrectly specified equations reveal that they imply quite different welfare values. Logit and probit models with broad specifications may be perfectly reasonable methods of estimating empirical relationships. Conditional utility

[5] *See, for example, Hanemann (1984) where he makes this assumption explicit in his [6]. As he discusses in his footnote 6, the more general formulation is the one adopted in this paper.*

models can be used to measure welfare from these estimates, but only with very rigid specifications for these models. In general, for welfare analysis of discrete choice, the unconditional utility framework is far more promising than conditional utility models.

Literature Cited

Bishop, R.C.; Heberlein, T.A. 1979. Measuring values of extramarket goods: Are indirect measures biased? American Journal of Agricultural Economics. 61: 926–30.

Bishop, R.C.; Heberlein, T.A.; Kealy, M.J. 1983. Contingent valuation of environmental assets: Comparisons with a simulated market. Natural Resource Journal. 23: 619–633.

Bowker, J.M.; Stoll, J.R. 1981. Use of dichotomous choice nonmarket methods to value the whooping crane resource. American Journal of Agricultural Economics. 63: 372–381.

Boyle, Kevin J.; Bishop, Richard C. 1985. The total value of wildlife resources: Conceptual and empirical issues. In: Proceedings, the AERE Workshop on Recreation Demand Modeling. May 17–18, 1985, Boulder, CO. Washington, DC: The Association of Environmental and Resource Economics and The Environmental Protection Agency.

Burt, O.; Brewer, D. 1971. Estimation of net social benefits for outdoor recreation. Econometrica. 39: 813–28.

Caulkins, P.; Bishop, R.; Bouwes, N. 1986. The travel cost model for lake recreation: A comparison of two methods for incorporating site quality and substitution effects. American Journal of Agricultural Economics. 68: 291–297.

Cichetti, C.; Fisher, A.; Smith, V.K. 1976. An econometric valuation of a generalized consumer surplus measure: The mineral king controversy. Econometrica. 44: 1259–76.

Hanemann, W.M. 1984. Welfare evaluation in contingent valuation experiments with discrete responses. American Journal of Agricultural Economics. 66: 332–41.

Kling, C. 1987. A simulation approach to comparing multiple site recreation demand models using Chesapeake Bay survey data. Journal of Marine Resource Economics. 4: 95–109.

Loehman, E.; De, V.H. 1982. Application of stochastic choice modeling to policy analysis of public goods: A case study of air quality improvements. Review of Economics and Statistics. 64: 474–80.

McFadden, D. 1981. Structural analysis of discrete data with econometric applications. In: Manski, C.; McFadden, D., eds. Econometric Models of Probabilistic Choice, Chapter 5. Cambridge, MA: MIT Press: 198–272.

Morey, E.; Rowe, R. 1985. The logit model and exact expected consumer's surplus measures: Valuing marine recreational fishing. In: Proceedings, the AERE Workshop on recreation Demand Modeling. May 17–18, 1985, Boulder, CO. Washington, DC: The Association of Environmental and Resource Economics and The Environmental Protection Agency.

Sellar, C.; Stoll, J.R.; Chavas, J.P. 1985. Validation of empirical measures of welfare change: A comparison of nonmarket techniques. Land Economics. 61: 156–75.

Sellar, C.; Stoll, J.R.; Chavas, J.P. 1986. Specification of the logit model: The case of the valuation of nonmarket goods. Journal of Environmental Economics and Management. 13: 382–90.

Small, K.; Rosen, H. 1981. Applied welfare economics with discrete choice models. Econometrica. 49: 105–30.

Smith, V.K.; Kaoru, Y. 1986. Modeling recreation demand within a random utility framework. Economic Letters. 22: 395–99.